Humboldt-Universitt zu Berlin

Master Thesis

submitted in partial fullfillment of the requirements for the degree of Master of Science in Statistics

Comparing Covariance Measures: A Systemic Risk Perspective

Written at the Ladislaus von Bortkiewicz Chair of Statistics

Author:

Christoph Hoffmann

(547425)

Supervisors:

Prof. Dr. Ostap Okhrin

Prof. Dr. Wolfgang K. Härdle

September 9, 2015

Contents

List of Figures

List of Tables

Abstract

In this paper I address the issue of predicting Conditional Value at Risk (CoVaR) by using a multivariate HEAVY and GARCH model and different realized covariance measures. Systemic risk forecasts are produced by Filtered Historical Simulation and conditional normality in the return distribution. The performance of the CoVaR is evaluated for all models, indicating the failure to describe systemic risk adequately during the financial crisis.

Keywords: CoVaR, Multivariate HEAVY, Systemic Risk, Realized Covariance

1 Introduction

Financial asset-return volatilities and correlations play a crucial role in several topics of financial theory and practice, including option pricing, asset allocation and risk management. With respect to the latter, Andersen et al. (2012) declare the measurement of risk and for that purpose the measurement of covariances as the key component of risk management. In this paper I provide a risk management framework to assess different ways to model and forecast covariances with the help of high-frequency covariance measures, thereby touching upon several of the most important research questions in financial econometrics. Not knowing the underlying latent covariance process, a practical backtesting procedure is adopted to evaluate the performance of different dynamic covariance models flavored by different high frequency measures. The multivariate risk measure Conditional Value at Risk ($CoVaR$) is introduced to enable model comparisons in a multivariate setting. The three research strands of 'Multivariate Covariance Modeling', 'High Frequency Covariance Measures' and 'Conditional Value at Risk' are put up for discussion: Which frequency measure is preferably used in forecasting conditional covariances? Do models that integrate high frequency measures outperform models that do not? Which distributional specifications lead to adequate systemic risk measures? The paper contributes to the current debate by presenting a multivariate framework for testing conditional covariance models and using the new class of multivariate HEAVY models to estimate the systemic risk measure $CoVaR$.

The recent prevalence of financial crises has fueled the search for systemic risk measures. One influential proposal by Adrian and Brunnermeier (2011) measures the tail distribution of the financial sector if a financial institution is at risk. This measure, called Conditional Value at Risk because it is conditioned on a specific event, enables an assessment of the risk this institution contributes to the financial sector. The basic idea of $CoVaR$ is illustrated in Figure 1: Different assets have different risk profiles which can be captured by their Value at Risk, which is depicted on the horizontal axis. But an isolated perspective on the risk profile might be misleading. Due to correlations between assets the risk profile will not only be determined by inherent factors but also in relation to the risk profile of other assets. The vertical axis presents the Value at Risk given that another asset is under stress (in this case the Bank of America). The three companies Du Pont, American Express and JP Morgan form an illustrative case: Even though they all have a similar *VaR* the risk profile changes if the Bank of America is below its *VaR*. In this case the conditional *VaR* for the stocks from the financial sector increases relative to the risk of Du Pont. The positive covariance between institutions from the financial sector drives up the conditional risk.

Forecasts of conditional risk measures require forecasts of conditional covariances. One model that allows for multi-period forecasts and incorporates high frequency measures is the multivariate HEAVY model by Noureldin et al. (2012). In the following sections I will develop the framework for conditional risk measurement and use the multivariate HEAVY and a multivariate GARCH model to forecast the Conditional Value at Risk. The *CoVaR* forecasts are produced by assuming a conditional distribution for both models and by Filtered Historical Simulation. The risk measure is then used to assess the performance of different covariance measures which are used in the multivariate HEAVY model.

2 CoVaR Framework for the Comparison of Covariance Measures

My approach to modeling *CoVaR* is a multivariate extension to the *VaR* forecasting framework proposed by Brownlees and Gallo (2009). At time t, the N-dimensional daily (close-to-close) return vector R_t is defined as

$$R_t = \Omega_t^{1/2} Z_t \quad Z_t \sim F, \tag{1}$$

where Ω_t is the $N \times N$ conditional covariance matrix of daily returns at time t and Z_t is an i.i.d. shock matrix which follows the cumulative distribution F. The one-day-ahead $100(1-p)\%$ *CoVaR*, to be defined in Section 4, gives the maximum one-day ahead loss of asset R_{it} conditional on some event in R_t:

$$CoVaR^p_{t|t-1} = G^{-1}(p)\, \Omega_t^{1/2}. \tag{2}$$

Here, $G^{-1}(p)$ denotes the inverse of the conditional distribution of asset R_{it} given the respective event in R_t. Since $G^{-1}(p)$ does not necessarily exist, simulation techniques

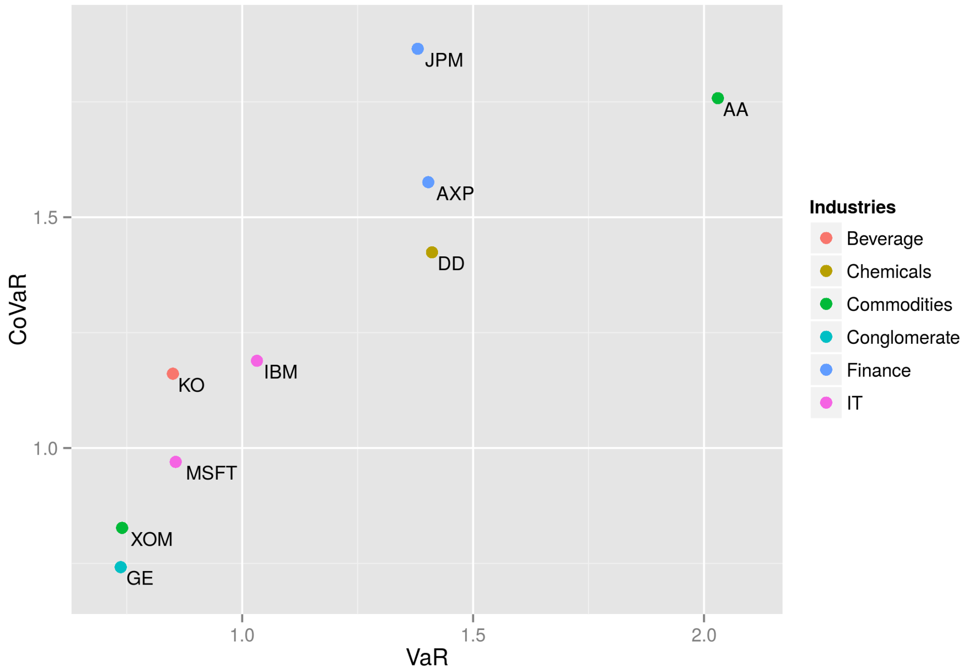

Figure 1: VaR vs. $CoVaR$: Simulated figures for the 31/12/2009 based on multivariate normality and daily realized covariance for the 9 daily stock returns of JP Morgan (JPM), International Business Machines (IBM), Microsoft (MSFT), Exxon Mobil (XOM), Alcoa (AA), American Express (AXP), Du Pont (DD), General Electric (GE) and Coca Cola (KO). Conditioning for $CoVaR_{5\%}$ is based on the Bank of America (BAC) beiing below its $VaR_{5\%}$.

are employed to determine the desired quantile. In Equation (2) Ω_t is fully determined by the information available at time $t-1$. Instead of following a standard GARCH procedure to model and forecast the conditional covariance matrix, a series of high-frequency covariance proxies (or 'covariance measures') is used to improve the forecasting accuracy. Such a proxy $RC_{(m,\delta)t}$ follows definition m, uses intra-daily data sampled at frequency δ and its expectation conditional on the information at time $t-1$ is denoted by $RC_{(m,\delta)t|t-1}$. The following sections discuss the fundamental building blocks of the *CoVaR* framework: The definitions of the covariance measures, the specification of *CoVaR*, the dynamics of conditional covariance matrix and tools to evaluate the obtained forecasts.

3 Realized Variances and Covariances

Early works by Merton (1980) and Nelson (1992) adverted the use of intraday data to develop improved volatility models. The starting point for the concept of realized variance in a univariate setting is the assumption that the increments of the logarithmic price p_t follow an Itō drift-diffusion process with drift coefficient $\mu(t)$ and instantaneous variance $\sigma(t)$:

$$\mathrm{d}\,p(t) = \mu(t)\,\mathrm{d}(t) + \sigma(t)\,\mathrm{d}\,W(t), \tag{3}$$

where $W(t)$ denotes a Wiener process. Assuming that the returns do not follow a drift I set $\mu(t)$ to zero. Realized volatility is then formally defined as

$$RV_{(v,\delta)t} = \sum_{i=2}^{n(\delta)} \left(p_{i,t} - p_{i-1,t}\right)^2. \tag{4}$$

Here δ denotes the sampling frequency for the ith intra day log price $p_{i,t}$ at date t. The summation index runs till $n(\delta) = n_{sec}/\delta$, where n_{sec} is the number of seconds in a trading day. The theory of quadratic variation shows that $RV_{(v,\delta)t}$ converges to the latent volatility as the sampling frequency increases, i.e. $\delta \to 0$ (see Andersen et al. (2001)). The key insight of this result is that summing sufficiently fine sampled squared intraday returns leads to (model free) ex-post realized volatility measures which render the true ex-post volatility observable.

Since the actual price process is not continuous a discrete sampling frequency is used to calculate the volatility measure in practice. One major debate in the literature on high frequency data analysis is the specification of this sampling frequency δ (see for example Aït-Sahalia et al. (2005), Hansen and Lunde (2006)). The convergence result suggests to use intraday returns based on the highest possible frequency but those returns suffer from market effects usually termed market microstructure noise, such as bid ask bounces, differences in trade sizes and their market impact or asymmetric information. A highly frequent variance measure would for instance not only capture the fluctuations of the price but also the price movement do to executing limit orders at the bid-ask spread. So even though the volatility of the underlying asset could be constant, these bounces would still be reflected in the measure. In a simple theoretical framework these effects would add up to the 'efficient' price in forming the observed one.

Sampling in a high frequency could then lead to a biased estimator which measures the variance of the noise and price process together. Possible remedies include the explicit integration of an error term in the price process, as in the two-scales realized volatility estimator by Zhang et al. (2005), or the deliberate choice to sample sparsely, at the expense of informational losses. A second concern in the construction of variance measures is the presence of jumps, i.e. large price movements over short time intervals. The presence of jumps, for instance due to macroeconomic news announcements or stock idiosyncrasies, disrupts the presented convergence results as well and needs to be accounted for, see e.g. Andersen et al. (2007). Several jump resistant volatility estimators, like the bipower variation estimator of Barndorff-Nielsen and Shephard (2004) have been proposed to maintain consistency in the case of volatility jumps.

The univariate setting can be generalized by introducing the multivariate Itō process

$$\mathrm{d}\,P(t) = M(t)\,\mathrm{d}(t) + \Omega(t)^{1/2}\,\mathrm{d}\,W(t). \tag{5}$$

In analogy to the univariate case, the increments of the log price $(N \times 1)$ vector P_t are described by the $(N \times 1)$ drift vector $M(t)$ and the 'square-root' of the instantaneous covariance matrix $\Omega(t)$. The realized covariance has become the benchmark covariance measure and is defined as

$$RC_{(rcv,\delta)t} = \sum_{i=1}^{n(\delta)} R_{it} R_{it}^{\top}, \tag{6}$$

where $n(\delta)$ is defined as above and R_{it} is the return vector at sampling period i.[1] In the absence of market microstructure noise, $RC_{rcv,\delta}$ converges to the integrated covariance matrix if δ converges to zero. In addition to the thread of market microstructure frictions, the multivariate estimators need to account for non-synchronous trading, see for instance Hayashi et al. (2005) or Voev and Lunde (2007), and positive definiteness of the covariance matrix. The multivariate realized kernel estimator by Barndorff-Nielsen et al. (2011) is an attempt to address these considerations.

[1] Since I mostly use $RC_{rcv,\delta}$ for different δ I will abbreviate the measure from now on to RCV_{δ}

4 Systematic Risk, Fragility and CoVaR

Systemic Risk has emerged as the focal point of current financial crisis research. The term financial crisis coins situations where the whole financial system is under stress and systemic events occur when stress in one financial institution or market leads in a sequential fashion to stress in another institution or market (De Bandt and Hartmann, 2000). Systemic risk can then be understood as the risk of experiencing such a systemic event. The definition of systemic risk is still an ongoing debate (Kaufman et al., 2000), but the most recent definitions focus on potential spillover effects and interdependencies between institutions and markets. This emphasis can be traced back to the increasing number of financial crises in recent financial history driven by spillovers between financial actors. A salient example is the 2007-09 financial crisis, which originated in the bursting of the housing market and cumulated in the (near) failure of several financial institutions, such as Lehman Brothers, and severe stock market losses (Brunnermeier, 2008). Adrian and Brunnermeier (2011) develop the concept of conditional Value at Risk (*CoVaR*) to measure systemic risk. Their particular focus lies on the measurement of an increased tail comovement of institutions' assets and liabilities in financial distress to assess the potential for systemic risk. The *CoVaR* is an extension of the Value at Risk (*VaR*) which only gauges the risk of an institution in isolation. As such, the authors define conditional Value at Risk in the following way:

> "We focus primarily on CoVaR, where institution is CoVaR relative to the system is defined as the VaR of the whole financial sector conditional on institution i being in a particular state, such as distress or the median state."
> (Adrian and Brunnermeier, 2011, p.3)

The marginal contribution to systemic risk of an institution is then defined as the difference between the *VaR* of the system given the institution is in distress and the *VaR* if it is in a 'normal' state. This difference is expressed as $\Delta CoVaR$. The obvious consequence of systemic risk management is to differentiate the regulatory treatment of institutions based on $\Delta CoVaR$ to internalize the costs of excessive (systemic) risk taking. Moreover Adrian and Brunnermeier (2011) indicate the generality of the risk measure CoVaR in the sense that it can also assess the risk spillovers from institution to institution. Applying $\Delta CoVaR^{j|i}$ to institutions i and j would gauge the increased risk of institution j given the change of institution's i state from 'normal' to stress. Interchanging the role of the financial sector and the institution offers a third perspective: Now the fragility of institution j is assessed by comparing its *VaR* conditional on the occurrence of a systemic event in the financial sector.

This paper applies the *CoVaR* framework to stock returns instead of institution and sector wide *VaR*'s.[2] The rational behind this shift in perspective is twofold:

From a systemic risk perspective it is still instructive to assess the sensitivity of market wide effects induced by firm-specific shocks (and vice versa) approximated by the comovement of market and company returns. This approach is followed by Acharya et al. (2012), who estimate the expected return of institution j during a systemic event on stock market i (i.e. a fall of 40%) and use this information to assess the additional need of capital to offset the loss in equity. The short forecasting horizon adopted in the following paragraphs would additionally serve a market risk perspective. Accounting for the comovement of asset or portfolio returns with market or separate asset returns could constitute an integral part in a wider stress testing framework and help to develop more comprehensive risk reports. The flexibility of the *CoVaR* framework needs to be taken carefully: As a pure measure of association which simply gauges conditional tail movements it cannot determine the causal direction between stressed events.

5 CoVaR Methodology

The classical formulation of value at risk defines VaR_q^i implicitly as the q% quantile,

$$\Pr(R_i \leq VaR_q^i) = q\%, \tag{7}$$

where VaR_q^i is expressed in terms of R_i, the loss of asset return i. The *CoVaR* extends the classical VaR definition by adding a conditional event on which this *VaR* the based. Following closely the initial formulation by Adrian and Brunnermeier (2011), I let $CoVaR_p^{j|\mathbb{C}(R_i)}$ denote the *VaR* of asset j conditional on some event $\mathbb{C}(R_i)$ of asset i. Therefore, $CoVaR_p^{j|\mathbb{C}(R_i)}$ is implicitly defined by the p% -quantile of the conditional

[2] The original *CoVaR* estimation exercise was based on public information of market-valued total assets of financial institutions.

probability distribution:

$$\Pr\left(R_j|\mathbb{C}(R_i) \leq CoVaR_p^{j|\mathbb{C}(R_i)}\right) = p\%. \tag{8}$$

A number of conditioning events are feasible. For instance, the definition of $\Delta CoVar$ involves the differences in *VaR* based on the conditioning events $\{R_i = VaR_q^i\}$ and $\{R_i = Median^i\}$.

6 Volatility and Covariance Modeling

6.1 Dynamic Modeling of Univariate Volatility

The conditional expectation and variance of daily univariate stock returns r_t given $\mathcal{F}_{t-1}$ are commonly defined as

$$\mathbb{E}[r_t|\mathcal{F}_{t-1}] = \mu_t \quad \text{and} \tag{9}$$

$$\text{VAR}(r_t|\mathcal{F}_{t-1}) = \sigma_t^2 = \mathbb{E}[(r_t - \mu_t)^2)|\mathcal{F}_{t-1}], \tag{10}$$

where $\mathcal{F}_{t-1}$ defines the information set available at time $t-1$. For reasons of simplicity I will restrain from specifying a mean equation and concentrate on the modeling of the conditional variance.

30 years ago, Engle (1982) and Bollerslev (1986) developed the GARCH class of models which proved successful in measuring and forecasting conditional volatilities. In its most simplified version, the conditional variance of the stock return r_t is specified as a linear function of its last period 'sample' and conditional variance r_{t-1}^2 and σ_{t-1}^2:

$$\text{VAR}(r_t|\mathcal{F}_{t-1}) = \sigma_t^2 = \alpha_0 + \alpha r_{t-1}^2 + \beta\sigma_{t-1}^2. \tag{11}$$

An alternative representation of Equation (11), obtained by recursive substitution, shows the close relationship between GARCH and exponential smoothing of squared returns:

$$\sigma_t^2 = \frac{\alpha_0}{1-\beta} + \alpha\sum_{j=1}^{\infty}\beta^{j-1}r_{t-j}^2. \tag{12}$$

The increased availability of high-frequency-based RV measures has shifted the focus from classical GARCH modeling to the integration of RV measures into the volatility model. The difference between a GARCH model, such as the one depicted in Equation (12), and dynamic-RV (DRV) models lies in the way the variances are inferred: GARCH style models solely use past daily returns and a specific model structure whereas DRV models employ (possibly noisy) ex post observations of true volatility.

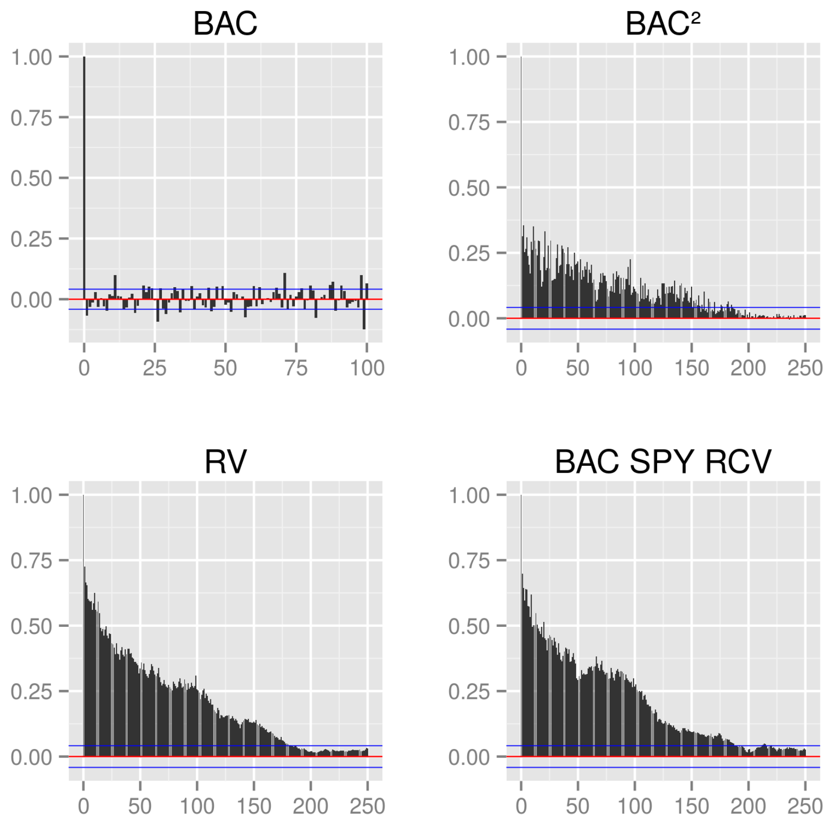

Figure 2: Sample autocorrelations. Panel (a) shows autocorrelations of daily returns of BAC (lag order between 1 and 100 days). Panel (b) displays the autocorrelation function for the squared series. (c) and (d) give autocorrelations for daily realized variance (5 min) and realized covariance of BAC and SPY respectively. The blue horizontal bands denote 95% Bartlett bands.

Figure 2 illustrates this point effectively. Even though the series r_t is serially uncorrelated, it is still a positively dependent series (compare Panel (a) and (b) respectively). This dependency is captured by the series r_t^2. Neglecting μ_t in equation 9 immediately shows that this series can be exploited to forecast conditional variances, dependent on a specific structure of the expectation function. Alternatively, panel (c) and (d) indicate a positive and highly statistically significant autocorrelation of realized covariances. If realized covariances are good proxies of conditional variances, DRV models may utilize their serial dependency to forecast conditional variances as well.

A natural starting point for the construction of DRV models are simple ARMA specifications for the process of realized variance. These models would use lagged values of RV measures and a weighted average of past errors to capture the dynamics observed in panel (c) of Figure 2. But the observed slow decay of the ACF is at odds with the exponential decay produced by ARMA type models. This would favor long-memory processes instead.

Two possibilities for long-memory processes are the ARFIMA and the HAR model depicted in Equations 13 and 14:

$$(1-L)^d RV_t = \alpha_0 + \upsilon_t \tag{13}$$

$$RV_t = \alpha_0 + \alpha_1 RV_{t-1} + \alpha_2 RV_{t-1}^{(w)} + \alpha_3 RV_{t-1}^{(m)} + \upsilon_t \tag{14}$$

. The ARFIMA process produces a hyperbolic decay of the ACF by means of fractional integration but is cumbersome to estimate. The estimation procedure is easier for the heterogeneous autoregressive (HAR-)model, proposed by Corsi (2009), which introduces different multi-period estimators of volatility. Here, weekly $RV_{t-1}^{(w)}$ and monthly $RV_{t-1}^{(m)}$ realized volatilizes are simply the normalized sums of one-period realized volatilities.

These richer dynamics are however only favorable for longer-term risk forecasts. Since I am only interested in short term forecasts i will not pursue long-memory dynamics. Instead I analyse the combination of GARCH and RV type models and their forecasting abilities.

6.2 Combining GARCH and RV-Modeling

One simple way of combining GARCH and RV based models is to add a RV measure as an additional regressor to the GARCH process displayed in Equation (11):

$$\sigma_t^2 = \alpha_0 + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma RV_{t-1}. \tag{15}$$

The resulting model class, discussed by Engle (2002b) and elaborated on by Lu (2005), is commonly referred to as GARCH-X.[3] In most empirical applications the RV measure drives out the ARCH coefficient α and the model reduces to

$$\sigma_t^2 = \alpha_0 + \beta\sigma_{t-1}^2 + \gamma RV_{t-1}. \tag{16}$$

Andersen et al. (2012) and Engle (2002b) explain this reduction by the superior qualities of RV measures as an estimator for true ex-post daily variation compared to squared daily returns and recommend the use of GARCH-X models instead of GARCH(1,1) when RV measures are available. For reasons of comparison a similar substitution exercise as in Equation (12) can be done:

$$\sigma_t^2 = \frac{\alpha_0}{1-\beta} + \gamma \sum_{j=1}^{\infty} \beta^{j-1} RV_{t-j}. \tag{17}$$

The GARCH-X process implies that volatility is an exponentially weighted moving average of past RV measures. Shephard and Sheppard (2010) state that in applied work, the weighting parameter β in GARCH-X is typically around 0.6-0.7 while in GARCH (see Equation (12)) it is around 0.91 or above. Hence GARCH-X is a weighted sum of very recent realized measures whereas GARCH possesses a longer memory and averages more data points. Inspecting the GARCH-X models in (15) and (16) closely reveals the necessity to integrate a dynamic process of the realized measure if multi-period volatilities are forecasted. This could be done for example by combining GARCH-X processes with the presented time series models for RV measures as in (13). Due to the fact that variation in realized measures is not accounted for, Hansen et al. (2012) frame GARCH-X models as incomplete and propose the class of Realized GARCH models as a way to complete the model:

$$\sigma_t^2 = \alpha_0 + \beta\sigma_{t-1}^2 + \gamma RV_{t-1}, \tag{18}$$

$$RV_{t-1} = \alpha_r + \beta_r\sigma_t^2 + \tau(z_t) + \upsilon_t. \tag{19}$$

Here, equation (19) contains the two error components $z_t \sim i.i.d.(0,1)$ and $\upsilon_t \sim i.i.d.(0,\sigma_\upsilon^2)$. Whereas (18) is already familiar, the so called measurement equation

[3] The X in the acronym GARCH-X signifies the treatment of the RV measure as an exogenous variable.

(19) describes the realized measure as a function of the current latent volatility plus a random innovation, which partly consists out of the leverage function $\tau(z_t)$. Since RV measures intraday volatility and r_t is a close-to-close return that spans 24 hours, Hansen et al. (2012) interpret the parameter β_r as the proportion of daily volatility which occurs during trading hours. The measurement equation 'completes' the model by specifying the dynamic properties of returns, i.e. $r_t = \sqrt{\sigma_t^2} z_t$, and realized variance. The relation between the return and the realized measure is modeled through the error term z_t. Respective substitutions reveal that the conditional variance in the Realized GARCH model follows an autoregressive process and realized volatility has an ARMA representation.[4] Within the class of multiplicative error models (MEM) Brownlees and Gallo (2009) define the conditional expectation of the realized measure as:[5]

$$\sigma_t^2 = \alpha_0 + \gamma\kappa_t, \tag{20}$$

$$\mathbb{E}[RV_t|\mathcal{F}_{t-1}] = \kappa_t = \alpha + \gamma RV_{t-1} + \beta\kappa_{t-1}. \tag{21}$$

The peculiar GARCH structure in (20) contains no lagged returns or conditional variances an utilizes the conditional expectation of the realized measure as the only explanatory variable. Shephard and Sheppard (2010) highlight that the smoothed version κ_t of lagged realized measures is used as an input variable for σ_t^2 even though the smoothing parameter β is chosen to optimize Equation (21) instead of (20). They prefer a structurally similar model with a raw version of the realized measure in the equation of conditional variance and an additional mean equation for RV measures which allow multi-period forecasts. This so called high-frequency- based volatility (HEAVY) model is defined as

$$\sigma_t^2 = \alpha_0 + \beta\sigma_{t-1}^2 + \gamma RV_{t-1}, \tag{22}$$

$$\mathbb{E}[RV_t|\mathcal{F}_{t-1}] = \kappa_t = \alpha_h + \gamma_h RV_{t-1} + \beta_h\kappa_{t-1}. \tag{23}$$

[4]This is why the authors label the model as a GARCH process (without an ARCH coefficient).

[5]The original formulation of the model contains an additional leverage effect which is suppressed to increase clarity.

In this two equation system the dynamics of the realized measure are treated differently as in (19): Similar to (22) the dynamics of the realized measure employ a GARCH like structure. The Realized GARCH however relates the realized measure back to the conditional variance.

6.3 Dynamic Modeling of Covariances

The following section extends the previous insights on volatility modeling to a multivariate setting. For this exercise I introduce a multivariate version of the univariate HEAVY model, a dynamic conditional correlation model and a simple multivariate GARCH model as a benchmark model. General surveys of multivariate volatility and GARCH models can be found in Chapter 10 of Tsay (2010) and Bauwens et al. (2006). A natural extension to the univariate conditional moment equations (9) and (10) is a multivariate return process with time-varying conditional mean and covariance:

$$R_t = M_t + \Omega_t^{1/2} Z_t, \quad Z_t \sim i.i.d.(0, \mathbb{I}_N), \tag{24}$$

$$\mathbb{E}[R_t|\mathcal{F}_{t-1}] = M_t = 0 \quad \mathbb{E}[R_t R_t^\top|\mathcal{F}_{t-1}] = \Omega_t. \tag{25}$$

Here, the $N \times N$ matrix $\Omega_t^{1/2}$ is a square-root representation of the covariance matrix Ω_t and $\mathbb{I}_N$ gives the identity matrix of the same dimensionality.[6] Conditional modeling of covariance matrices implies that Ω_t is a non-trivial function of the information set $\mathcal{F}_{t-1}$. Furthermore, the daily expected returns are assumed to be zero.[7] An alternative representation of (24) used later on is given by

$$P_t = R_t R_t^\top = \Omega_t^{1/2} \epsilon_t \Omega_t^{1/2} \quad \epsilon_t = Z_t Z_t^\top, \tag{26}$$

where Z_t is defined as in (24). A simple multivariate equivalent by Bollerslev et al. (1988) of the univariate GARCH in Equation (11) is

$$\mathrm{vech}(\Omega_t) = \mathrm{vech}(C) + B\,\mathrm{vech}(\Omega_{t-1}) + A\,\mathrm{vech}(R_{t-1} R_{t-1}^\top), \tag{27}$$

[6] For a positive definite matrix A the spectral decomposition $A = CDC^\top$, where D is the diagonal matrix of eigenvalues and C is the matrix of normalized eigenvectors, can be modified by taking the square roots of the eigenvalues to produce a square root matrix $A^{1/2} = CD^{1/2}C^\top$. This square root matrix serves as a square root of A since $A^{1/2}A^{1/2} = A$.

[7] The conditional mean vector M_t could alternatively be specified as a vectorial autoregressive moving average representation of R_t.

where the half-vectorization "vech" converts the lower triangular part of a symmetric matrix into a column vector of dimension $\frac{1}{2}N(N+1) \times 1$. Compatability requires B and A to be a $\frac{1}{2}N(N+1) \times \frac{1}{2}N(N+1)$ matrix. A general problem with multivariate systems is the fast growing number of parameters. Equation (27) requires the estimation of $N(N+1)(N(N+1)+1)/2$ parameters. One remedy to simplify the estimation process, introduced by Bollerslev et al. (1988), is to assume diagonality for B and A. As a consequence, the elements of Ω_t will only depend on their respective lagged values and (cross-)products of returns.

A different problem in multivariate modeling of covariances is the necessitated positivity of Ω_t. Engle and Kroner (1995) introduced the so called BEKK parametrization (named after Baba, Engle, Kraft and Kroner) to ensure positivity. The basic idea of the BEKK parametrization involves the pre- and postmultiplication of explanatory matrices with a parameter matrix, i.e. $BR_{t-1}R_{t-1}^\top B^\top = L_{t-1}$ to form a positive-definite inner product in the quadratic form $x^\top L_{t-1}x = y^\top y$ with $y = R_{t-1}^\top B^\top x$. The canonical representation of the GARCH BEKK(1,1) model in matrix notation is

$$\Omega_t = CC^\top + B\Omega_{t-1}B^\top + AR_{t-1}R_{t-1}A^\top. \tag{28}$$

The two alternative representations below illustrate the dynamics of the bivariate BEKK model if the parameter matrices A and B are assumed to be diagonal. In matrix notation this model becomes

$$\begin{aligned}
\begin{pmatrix} \sigma_{11,t} & \sigma_{12,t} \\ \sigma_{21,t} & \sigma_{22,t} \end{pmatrix} &= \begin{pmatrix} c_{11} & 0 \\ c_{21} & c_{22} \end{pmatrix} \begin{pmatrix} c_{11} & 0 \\ c_{21} & c_{22} \end{pmatrix}^\top \\
&+ \begin{pmatrix} b_{11} & 0 \\ 0 & b_{22} \end{pmatrix} \begin{pmatrix} \sigma_{11,t-1} & \sigma_{12,t-1} \\ \sigma_{21,t-1} & \sigma_{22,t-1} \end{pmatrix} \begin{pmatrix} b_{11} & 0 \\ 0 & b_{22} \end{pmatrix} \\
&+ \begin{pmatrix} a_{11} & 0 \\ 0 & a_{22} \end{pmatrix} \begin{pmatrix} r_{1,t-1}^2 & r_{1,t-1}r_{2,t-1} \\ r_{2,t-1}r_{1,t-1} & r_{2,t-1}^2 \end{pmatrix} \begin{pmatrix} a_{11} & 0 \\ 0 & a_{22} \end{pmatrix}.
\end{aligned} \tag{29}$$

Taking the vech operator after multiplying out the matrices shows immediately that the conditional variances and covariances depend only on their own respective lag and the (cross-)products of the errors:

$$\begin{pmatrix}\sigma_{11,t}\\ \sigma_{21,t}\\ \sigma_{22,t}\end{pmatrix} = \begin{pmatrix}c_{11}^2\\ c_{11}c_{21}\\ c_{21}^2+c_{22}^2\end{pmatrix} + \begin{pmatrix}b_{11}^2 & 0 & 0\\ 0 & b_{11}b_{22} & 0\\ 0 & 0 & b_{22}^2\end{pmatrix}\begin{pmatrix}\sigma_{11,t-1}\\ \sigma_{21,t-1}\\ \sigma_{22,t-1}\end{pmatrix} + \begin{pmatrix}a_{11}^2 & 0 & 0\\ 0 & a_{11}a_{22} & 0\\ 0 & 0 & a_{22}^2\end{pmatrix}\begin{pmatrix}r_{1,t-1}^2\\ r_{1,t-1}r_{2,t-1}\\ r_{2,t-1}^2\end{pmatrix}. \tag{30}$$

Equation (30) reveals that there are no interaction effects between the different elements of the conditional covariance matrix in the sense that $\frac{\partial\sigma_{ij,t}}{\partial\sigma_{kl,t-1}} \neq 0$ only if $i = k$ and $j = l$. Similarly, a change in previous returns on asset i does not effect the conditional variance of asset j.

Additional to the conditional variance Equation (25) the multivariate HEAVY model by Noureldin et al. (2012) adds, as in (23) a conditional mean equation for the the realized covariance measure:

$$\mathbb{E}[RC_t|\mathcal{F}_{t-1}] = K_t \tag{31}$$

The two equation system (31) and (25) together form the multivariate HEAVY model. Again, as in the univariate case, the basic idea is to exploit the information in the realized measure to forecast the conditional covariance of stock returns instead of or additional to using the outer product of returns. Noureldin et al. (2012) also adopt a BEKK parametrization to ensure positivity:

$$\Omega_t = CC^\top + B\Omega_{t-1}B^\top + ARC_{t-1}A^\top, \tag{32}$$

$$K_t = DD^\top + EK_{t-1}E^\top + FRC_{t-1}F^\top. \tag{33}$$

In its unrestricted version, the $(N \times N)$ matrices A, B, E and F consist of N^2 parameters, whereas the lower triangular $(N \times N)$ matrices C and D contain $N(N+1)/2$ parameters. It can be easily shown, by the argument applied above, that Ω_t and K_t are positive semidefinite if Ω_0 and K_0 are positive semidefinite as well.

Furthermore Ω_t and K_t will be positive definite if C and D are additionally of full rank. The dimensionality of the HEAVY model can be further reduced by enforcing a diagonalised or scalar form of A, B, E and F. In its diagonal form, the first equation of the bivariate HEAVY model reads as

$$\begin{aligned}\begin{pmatrix}\sigma_{11,t} & \sigma_{12,t}\\ \sigma_{21,t} & \sigma_{22,t}\end{pmatrix} &= \begin{pmatrix}c_{11} & 0\\ c_{21} & c_{22}\end{pmatrix}\begin{pmatrix}c_{11} & 0\\ c_{21} & c_{22}\end{pmatrix}^{\top} \\ &+ \begin{pmatrix}b_{11} & 0\\ 0 & b_{22}\end{pmatrix}\begin{pmatrix}\sigma_{11,t-1} & \sigma_{12,t-1}\\ \sigma_{21,t-1} & \sigma_{22,t-1}\end{pmatrix}\begin{pmatrix}b_{11} & 0\\ 0 & b_{22}\end{pmatrix} \\ &+ \begin{pmatrix}a_{11} & 0\\ 0 & a_{22}\end{pmatrix}\begin{pmatrix}rc_{11,t-1} & rc_{12,t-1}\\ rc_{21,t-1} & rc_{22,t-1}\end{pmatrix}\begin{pmatrix}a_{11} & 0\\ 0 & a_{22}\end{pmatrix}.\end{aligned} \tag{34}$$

The vech representation of this HEAVY model is depicted below:

$$\begin{pmatrix}\sigma_{11,t}\\ \sigma_{21,t}\\ \sigma_{22,t}\end{pmatrix} = \begin{pmatrix}c_{11}^2\\ c_{11}c_{21}\\ c_{21}^2 + c_{22}^2\end{pmatrix} + \begin{pmatrix}b_{11}^2 & 0 & 0\\ 0 & b_{11}b_{22} & 0\\ 0 & 0 & b_{22}^2\end{pmatrix}\begin{pmatrix}\sigma_{11,t-1}\\ \sigma_{21,t-1}\\ \sigma_{22,t-1}\end{pmatrix} \tag{35}$$

$$+ \begin{pmatrix}a_{11}^2 & 0 & 0\\ 0 & a_{11}a_{22} & 0\\ 0 & 0 & a_{22}^2\end{pmatrix}\begin{pmatrix}rc_{11,t-1}\\ rc_{21,t-1}\\ rc_{22,t-1}\end{pmatrix} \tag{36}$$

The models (29) and (34) are used in the empirical part of the paper to estimate and forecast the conditional covariance matrices required for forming the $CoVaR$ measure.

6.3.1 The Dynamic Conditional Correlation Model

Another alternative for modeling covariance matrices is given by decomposing the conditional covariance matrix into different matrices for conditional correlations and standard deviations:

$$\Omega_t = D_t \Gamma_t D_t. \tag{37}$$

Here Γ_t is the conditional correlation matrix which is pre- and post multiplied by the diagonal matrix of conditional standard deviations D_t. Different assumptions on the behavior of the correlation matrix allow the formulation of the so called Constant Conditional Correlation (CCC) and Dynamic Conditional Correlation (DCC) GARCH models (see Bollerslev (1990), Engle (2002a) and Tse and Tsui (2002)). A straightforward generalization of univariate modeling arises if the conditional correlations are assumed to be time-invariant, i.e. $\Gamma_t = \Gamma$. In this CCC-setup the conditional covariance is only driven by time-varying conditional variance which in turn can be estimated by the processes described in the previous sections. One advantage of the CCC model compared to the above multivariate GARCH model is the flexibility introduced by allowing to model conditional variances for different stocks with different model specifications. This increased flexibility is bought by suppressing the dynamic nature of covariance, which might only be acceptable for short term forecasting. One remedy, suggested by Engle (2002a) and Tse and Tsui (2002), treats the conditional correlations as time-varying within the class of GARCH models. Such a DCC model, assuming a GARCH(1,1) process for Γ_t, can be expressed as

$$Q_t = C + \beta Q_{t-1} + \alpha \left(e_{t-1} e_{t-1}^{\top} \right). \tag{38}$$

In this formulation e_t denotes the vector of standardized returns $e_t = R_t D_t^{-1}$ and Q_t is a normalized version of the conditional correlation matrix

$$Q_t = \operatorname{diag}\{Q_t\}^{1/2} \Gamma_t \operatorname{diag}\{Q_t\}^{1/2}, \tag{39}$$

which ensures an appropriate range for the individual correlations (see Andersen et al. (2012)). One drawback of the DCC model in this form is that the scalar structure of the parameters α and β implies identical dynamics for the conditional correlations.

7 Estimation and Inference

7.1 The Distribution of ϵ_t in the Multivariate HEAVY Model

The first equation of the HEAVY model in (26) is

$$P_t = \Omega_t^{1/2} \epsilon_t \Omega_t^{1/2}.$$

As in the original formulation of the model I choose the density of the innovation matrix ϵ_t to follow a Wishart distribution. This holds if the vector of daily returns has a multivariate normal distribution: $R_t = \Omega_t^{1/2} Z_t$ where $Z_t \overset{i.i.d.}{\sim} N(0, \mathbb{I}_N)$. The main argument for choosing a Wishart distribution is to restrict the support of P_t to be the space of positive semidefinite matrices (Noureldin et al. (2012)). It follows immediately from this distributional assumption and the characteristics of the Wishart distribution that P_t is conditionally Wishart distributed $P_t|\mathcal{F}_{t-1} \sim W_N(n, \Omega_t)$.[8] Due to the fact that the matrix P_t has rank 1 the authors of the HEAVY model recommend the usage of the singular Wishart density for the innovation matrix ϵ_t, i.e. $\epsilon_t \overset{i.i.d.}{\sim} SINGW_N(1, \mathbb{I}_N)$.[9]

7.2 Parameter Estimation

The HEAVY-P and HEAVY-V equations are exogenous in the sense that their respective model parameters are variation free (see Engle et al. (1983)). Thus I only use estimation properties of Noureldin et al. (2012) for the HEAVY-P equation. Let the HEAVY-P equation be parameterized with a finite-dimensional $(\delta \times 1)$ parameter vector $\theta \in \Theta \subset \mathbb{R}^\delta$. The log likelihood function $l_t(\theta)$ is defined for every observation $t \in T$ (see appendix A). Statistical inference is based on quasi-maximum likelihood estimation (QMLE) to accommodate for a potential misspecification of the likelihood function. The likelihood function to optimize is given by

$$l_t(\theta) = c - \frac{1}{2}(\log|\Omega_t| + \mathrm{tr}(\Omega_t^{-1} P_t)),$$

where all terms independent of θ are summarized into the constant c. Estimation requires an initialization of the process at $t = 0$ Therefore Ω_0 is assumed to be given and positive semidefinite .[10] The quasi-maximum likelihood estimator $\widehat{\theta}$ is the argument which maximizes the log-likelihood function L_t:

$$\widehat{\theta} = \arg\max_{\theta \in \Theta} L_T(\theta) = \arg\max_{\theta \in \Theta} \sum_{t=1}^{T} l_t(\theta).$$

The score vector S_t collects the $(1 \times \delta)$ vector of partial derivatives with respect to the parameter θ and equals

$$S_t(\theta) = \frac{\partial l(\theta)}{\partial \theta^\top} = 0.5[\mathrm{vec}(P_t)^\top - \mathrm{vec}(H_t)^\top](\Omega_t^{-1} \otimes \Omega_t^{-1})\frac{\partial \Omega_t}{\partial \theta^\top}.$$

[8] For the matrix S following a Wishart distribution $S \sim W_N(n, \Sigma)$ it holds that $ASA^\top \sim W_N(n, A\Sigma A^\top)$ for any nonsingular matrix A.

[9] The linear dependency in rows is immediately seen by comparing two arbitrary rows j and k of P_t, i.e. $R_{jt}(R_{1t}, R_{2t}, \cdots, R_{Nt})$ and $R_{kt}(R_{1t}, R_{2t}, \cdots, R_{Nt})$, where the one is just the other multiplied by a scalar. This holds if R_t contains at least one non-zero return.

[10] For reasons of feasibility I initialize all processes at the current respective matrix of realized covariances which therefore figures as a proxy for the conditional covariance matrix.

Noureldin et al. (2012) refer to the strong consistency results for the QMLE by Comte and Lieberman (2003) given that the model has a strictly stationary and ergodic solution. These results establish that $\widehat{\theta}$ is asymptotically normally distributed with standard errors given by the so-called "sandwich-estimator":

$$\sqrt{T}\left(\widehat{\theta}-\theta_0\right) \xrightarrow{d} N(0, \mathcal{I}^{-1}\mathcal{J}\mathcal{I}^{-1}),$$

$$\mathcal{J} = \mathbb{E}\left[S_t(\theta)^\top S_t(\theta)\right], \quad \mathcal{I} = -\mathbb{E}\left[\frac{\partial S_t(\theta)}{\partial \theta}\right].$$

The variance of $\widehat{\theta}$ is a scaled function of the outer product of the expected score pre- and post multiplied by the inverse of the Hessian.

The QMLE of the diagonal BEKK GARCH model and its parameter vector θ_B is similar to the BEKK HEAVY model and uses the fact that R_t follows a multivariate normal distribution. In this case the log likelihood function $l_{B,t}(\theta_B)$ for observation t is taken from Silvennoinen and Teräsvirta (2009) as

$$l_{B,t}(\theta_B) = \bar{c} - \frac{1}{2}\left(\log|\Omega_t| + R_t^\top \Omega_t^{-1} R_t\right).$$

7.3 CoVaR Estimation

The only difference between $CoVaR$ and VaR is the added conditioning set. Therefore the methodologies for forecasting $CoVaR$ are mostly drawn from the VaR apparatus. In their original contribution, Adrian and Brunnermeier (2011) estimate the $CoVaR$ based on a quantile regression and a bivariate normal GARCH(1,1) model. Several other extensions of univariate VaR measures to the multivariate $CoVaR$ setting are feasible.

GARCH or time-varying conditional covariance models allow for $CoVaR$ forecasts because they deliver one-day ahead conditional covariance forecasts and require a distributional assumption on the shock vector Z_t for estimating the parameters. Therefore it is possible to determine the one-day ahead return distribution and the quantiles of interest by scaling the presumed shock distribution with the conditional covariance forecast. Several parametric distributions for Z_t, like the multivariate normal or multivariate t distribution, have been applied to $CoVaR$ estimation.

The multivariate HEAVY model employs a bivariate normal shock vector which serves as a natural starting point for the following estimation exercise. In their GARCH set-up Adrian and Brunnermeier (2011) choose the conditioning set of asset return i to be $\{R_{it} = VaR_q^i\}$. This equalization facilitates the expression for the $CoVaR$ forecast because normality is preserved under conditioning, but it is rather unpractical from a backtesting perspective. The expectation of observing the conditioning set is zero by construction and testing forecasts by comparing them with empirical observations becomes unfeasible. I construct a richer conditioning set by including all observations below the VaR_q^i threshold: $\mathbb{C}(R_{it}) = \{R_{it} \leq VaR_q^i\}$. In order to estimate the CoVaR on the previous conditioning set I simulate a bivariate normal distribution with the forecasted conditional covariance matrix as an input parameter. The conditional return distribution based on time t thereby becomes $R_{t+1|t} \sim N(0, \Omega_{t+1|t})$. From the simulated returns I calculate the respective unconditional and conditional empirical $q\%$- and $p\%$-quantiles. It is worth noting that the returns are therefore conditionally normal which does not imply unconditional normality because the covariance dynamics inflate the tails, see Andersen et al. (2012) for a discussion on the caveats of normal return shocks. But these authors also emphasize that normal innovations only rarely provide an adequate description: Even Conditionally normality is sometimes not fat-tailed enough and asymmetric distributional features are impossible to replicate.

The potential pitfalls presented above can be avoided by Filtered Historical Simulation (FHS), see Diebold et al. (1998), Hull and White (1998) and Gurrola and Murphy (2015). The basic idea of FHS builds upon the previously estimated covariance matrices and their consistency while relaxing the (parametric) distributional assumption of the innovations, see Pritsker (2006). They are still treated as i.i.d. with zero mean and covariance matrix $\mathbb{I}$, but the distribution is only required to provide consistent estimates. If we treat the implied conditional covariances as correctly estimated, it is possible to identify the previous realizations of the return shocks as

$$\widehat{Z}_t = \widehat{\Omega}_t^{-1/2} R_t. \tag{40}$$

By sampling repeatedly with replacement from the series of filtered shocks $(\widehat{Z}_t)_{t=1}^{t=T}$ one can again determine the unconditional and conditional empirical quantiles.

8 Assessing the Failure Rate

Generally, the assessment of *VaR* forecasts focuses on the (un-)conditional coverage (adequacy) and on the precision (accuracy) of the forecasts. The first aspect asks the question on how often the *VaR*-threshold is breached and the second question is concerned with distance measures between the realized return and the breached *VaR*-threshold.

The *VaR* failure rate is the proportion of returns smaller than *VaR*. A number of tests have been proposed to assess the adequacy of *VaR* forecasts based on this failure rate, e.g. Christoffersen (1998) and Campbell (2005). The binary indicator of *VaR* failure gives rise to the so-called "hit" function:

$$I_{t+1}(q) = \begin{cases} 1 \quad \text{if} \quad R_{jt+1} \leq -VaR_{t+1|t}(q) \\ 0 \quad \text{if} \quad R_{jt+1} > -VaR_{t+1|t}(q) \end{cases} . \tag{41}$$

The sequence of hit functions $(I_{t+1}(q))_{t=1}^{t=T}$ indicates at which point in time the realized returns were below the forecasted *VaR*. Intuitively, the sum over the individual hit functions relative to T should be close to the adopted $q\%$-quantile. This is formalized as the unconditional coverage property, which states that the probability of observing a loss greater than VaR must be $q\%$. A deviation of the empirical failure rate in the sense that the losses exceed the *VaR* to often would indicate a systematic bias. An early LR test authored by Kupiec (1995) is based on the unconditional coverage property in stating the null hypothesis that the expectation of I_t equals $q\%$: $H_0 : \mathbb{E}(I_t) = q\%$. The test-statistic PF follows asymptotically a χ_1^2 distribution and equals

$$PF = 2 \log \left(\left(\frac{1-\widehat{q}}{1-q} \right)^{T-I(q)} \left(\frac{\widehat{q}}{q} \right)^{I(q)} \right) \overset{H_0}{\sim} \chi_1^2, \tag{42}$$

where $\widehat{q} = \frac{1}{T} I(q)$ is the empirical failure rate defined by $I(q) = \sum_1^T I_t(q)$.

Christoffersen (1998) extends this early unconditional coverage test by introducing an additional property of independence: Any two elements of the sequence $(I_{t+1}(q))_{t=1}^{t=T}$ should be independent from each other. A dependence of the terms implies a time dependent probability of *VaR* failure and an unconditional assessment of coverage would be misleading. One indication of dependence is a clustering of *VaR* failures. One

LR test of independence by Christoffersen (1998) is formally described in Appendix C of the paper.

The assessment of accuracy requires the definition of a loss function. The information of the hit function is limited because it disregards the difference between the realized return below the *VaR* threshold and the threshold itself. I follow the definition by Lopez et al. (1999) and integrate a quadratic magnitude term into the binomial loss function such that

$$L(VaR_{t+1|t}(q), R_{jt+1}) = \begin{cases} 1 + (R_{jt+1} - VaR_{t+1|t}(q))^2 & \text{if} \quad R^i_{t+1} \leq -VaR_{t+1|t}(q) \\ 0 & \text{if} \quad R_{jt+1} > -VaR_{t+1|t}(q) \end{cases}. \tag{43}$$

This function can provide additional information on the ability of the *VaR* model to forecast the lower tail of the return distribution. Tests based on the loss function do need additional information on the stochastic behavior of the model. In principle this approach is feasible, because we have simulated the tail distributions in forecasting the *CoVaR*. A comparison of the average realized loss function with the expected loss of our simulation exercise would lead to further insights into the lower tail behavior of the returns. At this point I restrain from statistical testing and will only report the sample average loss $\widehat{L}$:

$$\widehat{L} = \frac{1}{T}\sum_{1}^{T} L\left(VaR_{t+1|t}(q), R_{jt+1}\right). \tag{44}$$

In the empirical part of the paper I will extend these tests to the CoVaR scenario. In order to do so I redefine the hit-function to account for the conditioning set $\mathbb{C}(R_i)$:

$$\widetilde{I}^C_{t+1}(p) = \begin{cases} 1 & \text{if} \quad R_{jt+1} \leq -CoVaR_{t+1|t}(p) \mid R_{it+1} \in \mathbb{C}(R_i) \\ 0 & \text{if} \quad R_{jt+1} > -CoVaR_{t+1|t}(p) \mid R_{it+1} \in \mathbb{C}(R_i) \end{cases}. \tag{45}$$

The integration of the conditioning set, defined in Section 4, is necessary to evaluate the adequacy of the model. Otherwise, the results would be confounded by events the *CoVaR* does not intent to cover. One example relates back to figure 1: The *VaR* of Alcoa is higher than its *CoVaR* and therefore the returns will stabilize if the conditioning asset is under stress. An unconditional evaluation of CoVaR would lead to a misleadingly high $CoVaR$ failure rate and the (correct) model would be dismissed. The empirical failure rate $\widehat{p}$ is therefore redefined as

$$\widehat{p} = \frac{1}{\widetilde{\widetilde{T}}} \widetilde{I}(p), \quad \widetilde{I}(p) = \sum_{1}^{\widetilde{T}} \widetilde{I}(p), \tag{46}$$

where $\widetilde{T}$ is the cardinality of the realized conditioning set $\#\mathbb{C}(R_i t)$. One problem of the conditional testing setup for backtesting purposes is the reduced sample size of failures if the conditioning set (and p) is too small. If $T = 1000$ and p and q equal 5%, the expected number of $CoVaR$ failures would only equal 2.5. For the $CoVaR$ evaluation exercise I do not test independence because a simple test as in Appendix C would disregard the differences in time between VaR failures and thereby adopt an oversimplified concept of 'neighborhood'.

9 Empirical Application

9.1 Data and Stylized Facts

For the empirical exercise I utilize the high-frequency stock market data provided by Noureldin et al. (2012). They develop the multivariate HEAVY model to forecast the conditional covariance matrix of open-to-close daily returns for Spyder (SPY), a S&P 500 exchange traded fund, and Bank of America (BAC), a highly liquid stock in the Dow Jones Industrial Average (DJIA). The source of the data is the Trade and Quote database of the New York Stock Exchange (NYSE) and the sample spans a period from February 1, 2001 to December 31, 2009. Figure 3 provides a first description of the data set, showing the return and 5min realized covariance series for the entire sampling period. Panels (1) to (3) indicate the sharp increase in volatility linked to the financial market crisis in 2008-2009, whereas the rise in BAC volatility is stronger compared to SPY. The daily correlation between SPY and BAC, approximated by the 5min realized covariance measure, is almost always positive and slightly increasing in time with a number of prominent negative deviations. Additionally Panels (2) and (3) exhibit volatility clustering for the realized volatility measures, suggesting a GARCH like structure as in Equation (33) for the realized covariance process.

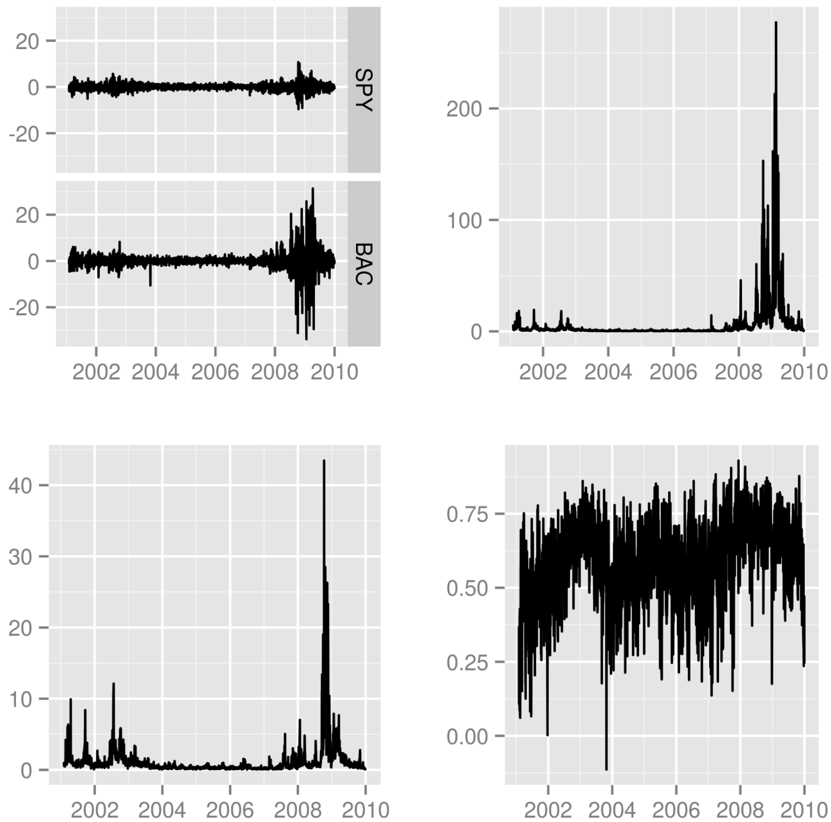

Figure 3: Time series plots for close-to-close daily returns and 5min realized covariances of BAC and SPY for the entire sample period. Panel 1 displays the daily return series. The top right and bottom left panel picture 5min realized variances for BAC and SPY respectively. The last panel shows realized correlations between BAC and SPY.

9.2 Realized Measures

The multivariate HEAVY model is estimated for a number of different sampling frequencies of $RC_{(rcv,\delta)}$ with δ equal to $1, 5, 10, 15$ and 30m. The multivariate realized kernel estimator by Barndorff-Nielsen et al. (2011) is introduced as an alternative measure for the HEAVY equation. Table 1 provides the means, standard deviations and first order autocorrelations of the integrated covariance measures. The moments show no trends or other salient differences between the realized variances of BAC and SPY or their realized covariances. The RCV_1 appears to have a slightly higher variance for SPY and BAC which could be an indication of prevalent market microstructure effects.

		RCV_1	RCV_5	RCV_{10}	RCV_{15}	RCV_{30}	RCV_K
V.SPY	Mean	1.19	1.12	1.09	1.07	1.06	1.28
	Std.Dev.	2.45	2.25	2.16	2.29	2.36	2.89
	$\widehat{\rho}_1$	0.66	0.68	0.68	0.61	0.60	0.64
V.BAC	Mean	6.06	5.46	5.36	5.25	5.41	7.27
	Std.Dev.	21.08	16.81	17.34	16.73	18.33	24.67
	$\widehat{\rho}_1$	0.63	0.73	0.66	0.68	0.63	0.68
Cov.S.B	Mean	1.36	1.46	1.47	1.46	1.49	1.75
	Std.Dev.	3.75	3.58	3.49	3.8	4.03	4.53
	$\widehat{\rho}_1$	0.7	0.7	0.7	0.64	0.64	0.7

Table 1: Descriptive statistics for the different covariance measures used in the empirical application. $\widehat{\rho}_1$ provides an estimate of lag one autocorrelation.

9.3 Estimation of HEAVY Model and Residual Diagnostics

The first step of the CoVaR framework consists in estimating the HEAVY model for the different realized covariance measures. Estimated parameters for the entire sample are presented in Table (2) alongside the estimates for the benchmark BEKK GARCH model.

The estimated parameters of RCV_δ are all similar to each other and converge slowly towards the GARCH parameters as δ increases. This was to be expected since the GARCH process constitutes the limiting process obtained by reducing the frequency of intraday returns sampled. Compared to the modeling results of Noureldin et al. (2012) for open-to-close returns the parameter values for GARCH are quite similar whereas the $\widehat{a}_1$ and $\widehat{a}_2$ parameters for the HEAVY model appear to be significantly higher in the close-to-close setting. Since the most important difference between these series is the added overnight volatility, the up scaling of the realized covariance parameters could reflect the increased (relative) relevance of realized covariance for forecasting the

Parameters	RCV_1	RCV_5	RCV_{10}	RCV_{15}	RCV_{30}	RCV_K	GARCH
$\widehat{a}_1$	0.714	0.712	0.723	0.688	0.555	0.756	0.254
$\widehat{a}_2$	0.779	0.777	0.766	0.787	0.682	0.778	0.289
Var(SPY)	0.509	0.507	0.523	0.474	0.308	0.571	0.064
Var(BAC)	0.607	0.604	0.587	0.620	0.465	0.605	0.083
Cov	0.556	0.553	0.554	0.542	0.378	0.588	0.074
$\widehat{b}_1$	0.742	0.796	0.801	0.833	0.893	0.741	0.965
$\widehat{b}_2$	0.717	0.775	0.796	0.8	0.846	0.713	0.959
Var(SPY)	0.55	0.634	0.641	0.695	0.798	0.548	0.931
Var(BAC)	0.514	0.601	0.633	0.64	0716	0.508	0.92
Cov	0.532	0.617	0.637	0.667	0.756	0.528	0.925

Table 2: Diagonal HEAVY and GARCH estimation results for SPY and BAC using different realized covariances. Var(SPY),Var(BAC) and Cov give respective parameters for the vech representation.

conditional covariance matrix.

An evaluation of the fitted model can be based on residual diagnostics. The distributional assumptions of section 7.1 imply that the shock vector Z_t is multivariate normal with zero expectation and covariance matrix $\mathbb{I}_n$. Standardization of the returns $R_t = \Omega^{1/2} Z_t$ gives rise to the residuals $\widehat{Z}_t = \widehat{\Omega}_t^{-1/2} R_t$ of fitting the model, where $\widehat{\Omega}_t$ is the implied conditional covariance of the fit. Figure (4) displays the residuals for SPY, BAC and their covariances using a HEAVY model with a 5min realized covariance measure.[11] The residuals of BAC and SPY appear to be independent with zero expectation, but still contain a high degree of variation. The covariance residuals variate around zero and exhibit an asymmetric pattern with a higher quantity of large negative deviations. Compared to the HEAVY model for open-to-close returns, see Noureldin et al. (2012), the close-to-close model is less capable of describing the return distribution. This impression is amplified by figure 5 which shows a Q-Q plot of the BAC residuals because they have a marginal normal distribution by assumption. A Mardia's

[11]The following figures are all based on a 5min realized volatility measure. 5 deviant (covariance) residuals were removed from Figure 4 to improve visibility.

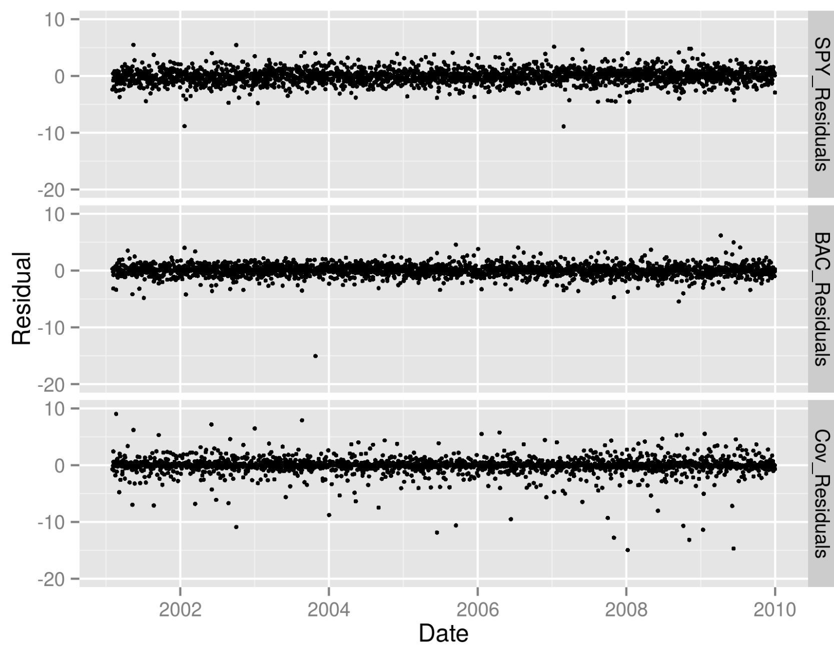

Figure 4: Implied Residuals $\widehat{Z}_t$ of the 5min HEAVY model.

test of multivariate normality clearly rejects the null hypothesis that the returns are normally distributed ($p_{\widehat{S}} < 0.0001$, $p_{\widehat{K}} < 0.0001$).

9.4 Forecasting CoVaR

9.4.1 Procedure Based on Presumed Normal Distribution of Return Shocks

The multivariate HEAVY and GARCH model is used to forecast the one-day ahead conditional covariance matrix $\Omega_{t+1|t}$. For this purpose the models are estimated with a rolling window consisting of 900 days which then forms the information set of time t. The forecast is given by

$$\Omega_{t+1|t} = \widehat{C}\widehat{C}^{\top} + \widehat{B}\Omega_t\widehat{B}^{\top} + \widehat{A}RC_t\widehat{A}^t,$$

where $\widehat{C}, \widehat{B}$ and $\widehat{A}$ are the respective QML matrix estimates of C, B and A. This procedure is repeated by shifting the estimation window one day ahead until the last observation of the sample is estimated. Overall I collect a series of 1342 covariance forecasts. Exploiting the assumption of bivariate normal return shocks I forecast the

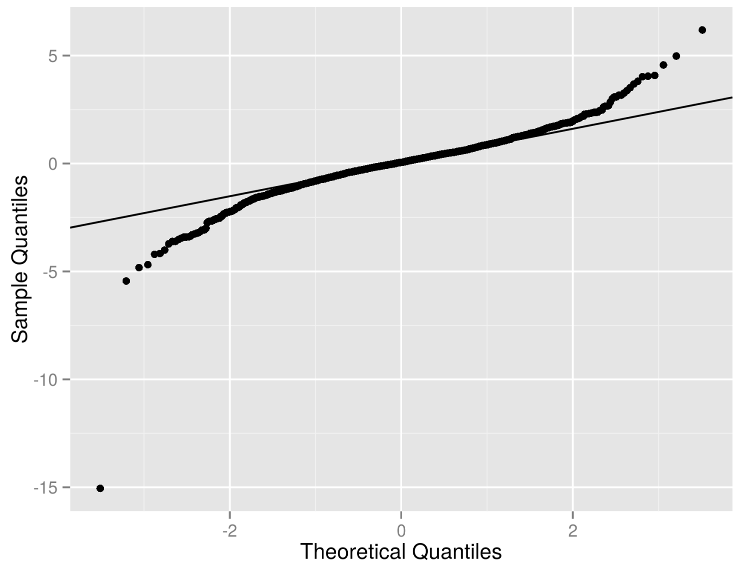

Figure 5: Normal Q-Q plot for the implied BAC residuals $\widehat{Z}_t$ from a 5min HEAVY model.

CoVaR of period $t+1$, $CoVaR^p_{t+1|t}$, by simulating one million draws from a bivariate normal distribution with mean zero and covariance matrix $\Omega_{t+1|t}$. The CoVaR is estimated as the empirical quantile of the simulated conditional distribution. As a conditioning set $\mathbb{C}(R_1 t+1)$ I choose the unconditional 35%-quantile of SPY. The final BAC-CoVaR forecast is based on a 95%-confidence level ($p = 5\%$). Figure 6 gives a first visual impression of the fit for the 5min HEAVY model. In principle the CoVaR forecast is able to peg the increasing and abating volatility of BAC returns during the financial crisis and its aftermath. Furthermore, the CoVaR exceeds the unconditional VaR and thereby accommodating the increased fragility induced by reduced stability. But Figure 6 also illustrates huge deficiencies of the CoVaR forecast. The indicator variable of CoVaR failure hints at a dependency of the unconditional conditional failure rate: CoVaR failures cluster heavily around the financial markets crisis in 2008-2009 indicating only poor adequacy of the fit. Additionally, the financial markets cluster exhibits severe excess returns which question the accuracy of the forecast as well.

A thorough evaluation of $CoVaR^p_{t+1|t}$'s predictive ability requires a preliminary examination of the conditioning set. This is necessary because a potential cumulation of SPY-$VaR_{0.035}$ failures could lead to an increased CoVaR failure rate. Figure 7 displays the daily close-to-close returns of SPY and the estimated $VaR_{0.035}$ and VaR

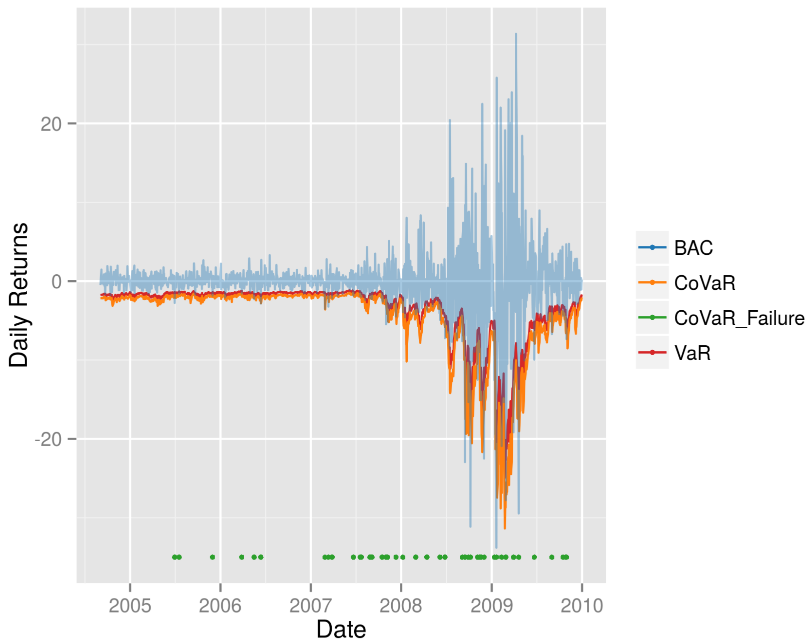

Figure 6: Estimated $CoVaR_{0.05}$ for daily returns of BAC based on the multivariate HEAVY model with a 5min covariance measure. The conditioning set is the $VaR_{0.35}$ of SPY. The displayed variable VaR shows the estimated unconditional $VaR_{0.05}$ of BAC. BAC reports the daily returns of BAC and CoVaR-Failure indicates timepoints where the conditional returns exceed the CoVaR.

failures. The visual examination is impeded since adequacy and accuracy are not as salient at the center of the distribution as they are in the tails. The failures of $VaR_{0.035}$ are not as clustered as the CoVaR failures with a slight increase in failures between mid 2007 and mid 2009. The excess failure is again more severe in this period of increased volatility. The evaluation measures for the CoVaR forecasting exercise under normality are collected in Table 9.4.1.

The forecasting abilities of the different covariance measures under conditionally normal distributed returns appear to be equal. The empirical failure rate of CoVaR and VaR are similar across the different specifications. All models fail the unconditional coverage tests for the forecasted VaR and CoVaR. The VaR forecasts underestimate the 35% quantiles and cover only 30% of the return distribution. The null hypothesis of equal transition probabilities for the VaR can not be rejected at the 5% significance level. Over the forecast horizon of 1342 periods it is expected that 23 CoVaR failures

Figure 7: Estimated $VaR_{0.35}$ for daily returns of SPY based on the multivariate HEAVY model with a 5min covariance measure. VaR-Failure indicates timepoints where the returns exceed the VaR.

$(= 1342 \times q \times p)$ occur given that the model specifications are correct. All models overestimate the CoVaR and exhibit a conditional failure rate of 9-10% instead of 5%. In accordance with the the residual diagnostics the distributional assumption about the returns prove to be misspecified. The lower tail of the distribution is too fat as to be adequately modeled by normality. This misspecification is so severe that the different performances of the covariance measures are of no importance. Given that all specifications fail, it is still interesting to notice that the classical GARCH model comes in first with respect to empirical failure of CoVaR and its Loss function. As an intermediate result one can state that, given the pictured situation, an adequate distributional specification is more important than flavoring GARCH type models with high-frequency data. In such a situation considerations about potential market microstructure effects in higher sampled covariance measures are beside the point. Figures 8 and 9 provide further insights into the adequacy and accuracy of the $CoVaR$ forecasts. The first plot of Figure 8 shows the difference between the conditional $CoVaR$ and the unconditional

	RCV_1	RCV_5	RCV_{10}	RCV_{15}	RCV_{30}	RCV_K	GARCH
Emp. failure VaR	0.31	0.31	0.31	0.30	0.30	0.31	0.3
PF-VaR	10	10.75	11.93	12.3	12.34	12.34	12.34
P-value	0.002	0.001	0.001	0.0004	0.0004	0.0004	0.0004
LR-I	2.31	2.23	2.31	2.08	2.65	1.79	1.09
P-value	0.13	0.14	0.13	0.15	0.10	0.18	0.3
Total failure CoVaR	42	43	40	38	40	40	36
Emp. failure CoVaR	0.101	0.104	0.098	0.093	0.098	0.097	0.091
PF-CoVaR	17.9	19.7	15.47	12.8	15.68	15.37	11.26
P-value	0	0	0	0.0004	0	0	0.0008
$\widehat{L}$	2.82	2.52	2.53	2.58	2.58	2. 47	1.77

Table 3: Evaluation results for VaR and $CoVaR$ under the distributional assumption of normality. P values below 0.0001 are denoted as 0.

VaR of BAC (green line). Both risk measures behave in unison until the height of the financial crisis. The gap between them increases which is probably related to the increase in covariance between SPY and BAC. In this case the *CoVaR* of BAC would rise relative to the unconditional *VaR*. This highlights the importance of conditional risk measures in times of financial stress. The second panel of Figure 8 provides a closer look at the performance of *CoVaR* during the financial crises. *CoVaR* failures occurred mostly and most severely around 2009. A display of the excess losses is given in Figure 9. This figure highlights two points: On the one hand, the excess losses in times of financial ease are fairly small and would favor the adoption of a *CoVaR* model based on conditional normality. On the other hand, the inaccuracy of the model is most severe in times of financial stress which is exactly the purpose it was designed for.

9.4.2 Procedure Based on Filtered Historical Simulation

Based on the poor findings for the multivariate normal distribution I decided to discard the model assumptions for the return equation and suggest that the data would be

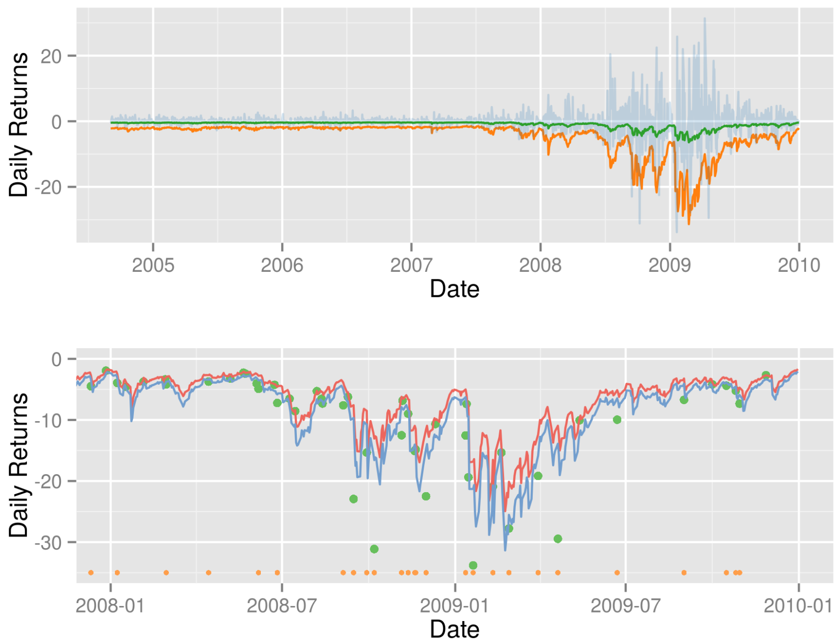

Figure 8: Panel 1 displays the *CoVaR* of BAC (orange line) and the difference between the 5% conditional and uncoditional *VaR* (green line). Panel 2 provides a close up of Figure 6 during the financial crisis. Estimates are based on the 5min HEAVY model.

better fitted by simulating the standardized returns. Similarly to the *CoVaR* estimation exercise beforehand I generated at each point in time a simulated sample of one 100.000 draws to determine the quantiles and conditional quantiles empirically. Table 9.4.2 gives the evaluation results for a selected variety of models. The FHS forecasting approach displays even worse results than the distributional assumption of conditional normality. The empirical failure rates for the VaR of SPY are slightly lower than the rates for conditional normality. The tests for unconditional coverage do therefore all reject the null hypothesis of adequacy. Only the GARCH model can not reject the null hypothesis of independent VaR failures. The performance of CoVaR is incredibly bad, estimating an average conditional quantile of 20% instead of the postulated conditional 5%-quantile.

Figure 9: Excess loss of CoVaR (green line) and CoVaR (orange line) for the 5min HEAVY model.

10 Conclusions

The aim of the paper was to differentiate between different models and high frequency measures by evaluating their respective performance in forecasting Conditional Value at Risk. This exercise was not feasible because all models failed to form adequate forecasts of Conditional Value at Risk. This holds true for estimations based on the distributional assumption of normal innovations and for the method of Filtered Historical Simulation. This holds also true for the simple GARCH model and the realized covariance flavored HEAVY model. One of the main results of the paper is that using realized covariances is unnecessary as long as the error distributions are not specified correctly. These distributional misspecifications outweigh potential moment misspecifications and correcting them should be of highest priority. Given the quite similar performances of the models under misspecification it is rather doubtful that one of the covariance measures emerges as the front runner in further investigations. A number of extensions to the presented risk framework are possible: Different distributional assumptions for the estimation and filtering of the conditional covariances can be made.

	RCV_1	RCV_5	RCV_{30}	GARCH
Emp. failure VaR	0.29	0.29	0.29	0.27
PF-VaR	25.4	21.4	20.89	35.2
P-value	0	0	0	0
LR-I	3.12	5.72	4.82	1.36
P-value	0.08	0.02	0.03	0.24
Total failure CoVaR	84	83	75	78
Emp. failure CoVaR	0.22	0.21	0.19	0.21
P-value	0	0	0	0
$\widehat{L}$	4.8	5.1	4.9	5.1

Table 4: Evaluation results for VaR and $CoVaR$ under Filtered Historical Simulation. P values below 0.0001 are denoted as 0.

Additionally, the number of periods to estimate the residuals for the FHS approach can be reduced. This would provide more shocks that are qualitatively similar to the one that is expected if the nature of the shock is changing. Given that a proper distribution is available, one can utilize the entire HEAVY model to form and evaluate multi-period forecasts of Conditional Value at Risk.

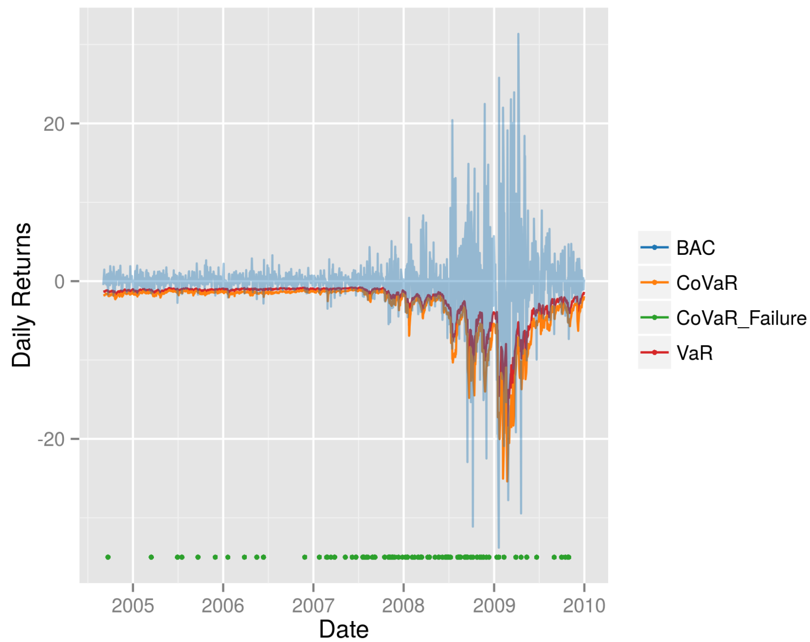

Figure 10: Estimated $CoVaR_{0.05}$ for daily returns of BAC based on the multivariate HEAVY model with a 5min covariance measure. The conditioning set is the $VaR_{0.35}$ of SPY. The displayed variable VaR shows the estimated unconditional $VaR_{0.05}$ of BAC. BAC reports the daily returns of BAC and CoVaR-Failure indicates timepoints where the conditional returns exceed the CoVaR. CoVaR Forecasts by Filtered Historical Simulation.

References

Acharya, V., Engle, R., and Richardson, M. (2012). Capital shortfall: A new approach to ranking and regulating systemic risks. *The American Economic Review*, pages 59–64.

Adrian, T. and Brunnermeier, M. K. (2011). Covar. Technical report, National Bureau of Economic Research.

Aït-Sahalia, Y., Mykland, P. A., and Zhang, L. (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *Review of Financial Studies*, 18(2):351–416.

Andersen, T. G., Bollerslev, T., Christoffersen, P. F., and Diebold, F. X. (2012). Financial risk measurement for financial risk management. Technical report, National Bureau of Economic Research.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics*, 61(1):43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Vega, C. (2007). Real-time price discovery in global stock, bond and foreign exchange markets. *Journal of International Economics*, 73(2):251–277.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2011). Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *Journal of Econometrics*, 162(2):149–169.

Barndorff-Nielsen, O. E. and Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2(1):1–37.

Bauwens, L., Laurent, S., and Rombouts, J. V. (2006). Multivariate garch models: a survey. *Journal of Applied Econometrics*, 21(1):79–109.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327.

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model. *The Review of Economics and Statistics*, pages 498–505.

Bollerslev, T., Engle, R. F., and Wooldridge, J. M. (1988). A capital asset pricing model with time-varying covariances. *The Journal of Political Economy*, pages 116–131.

Brownlees, C. T. and Gallo, G. M. (2009). Comparison of volatility measures: a risk management perspective. *Journal of Financial Econometrics*, page nbp009.

Brunnermeier, M. K. (2008). Deciphering the liquidity and credit crunch 2007-08. Technical report, National Bureau of Economic Research.

Campbell, S. D. (2005). A review of backtesting and backtesting procedures. Technical report, Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review*, pages 841–862.

Comte, F. and Lieberman, O. (2003). Asymptotic theory for multivariate garch processes. *Journal of Multivariate Analysis*, 84(1):61–84.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7(2).

De Bandt, O. and Hartmann, P. (2000). Systemic risk: a survey.

Diebold, F. X., Schuermann, T., and Stroughair, J. D. (1998). *Pitfalls and opportunities in the use of extreme value theory in risk management*. Springer.

Engle, R. (2002a). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3):339–350.

Engle, R. (2002b). New frontiers for arch models. *Journal of Applied Econometrics*, 17(5):425–446.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007.

Engle, R. F., Hendry, D. F., and Richard, J.-F. (1983). Exogeneity. *Econometrica: Journal of the Econometric Society*, pages 277–304.

Engle, R. F. and Kroner, K. F. (1995). Multivariate simultaneous generalized arch. *Econometric Theory*, 11(01):122–150.

Gurrola, P. and Murphy, D. (2015). Filtered historical simulation value-at-risk models and their competitors. Technical report.

Hansen, P. R., Huang, Z., and Shek, H. H. (2012). Realized garch: a joint model for returns and realized measures of volatility. *Journal of Applied Econometrics*, 27(6):877–906.

Hansen, P. R. and Lunde, A. (2006). Realized variance and market microstructure noise. *Journal of Business & Economic Statistics*, 24(2):127–161.

Hayashi, T., Yoshida, N., et al. (2005). On covariance estimation of non-synchronously observed diffusion processes. *Bernoulli*, 11(2):359–379.

Hull, J. and White, A. (1998). Incorporating volatility updating into the historical simulation method for value-at-risk. *Journal of Risk*, 1(1):5–19.

Kaufman, G. G. et al. (2000). Banking and currency crises and systemic risk: Lessons from recent events. *Federal Reserve Bank of Chicago Economic Perspectives*, 24(3):9–28.

Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives*, 3(2).

Lopez, J. A. et al. (1999). Methods for evaluating value-at-risk estimates. *Economic Review*, 2:3–17.

Lu, Y. (2005). *Modeling and forecasting daily stock return volatility with intra-day price fluctuation information*. PhD thesis, University of Pennsylvania.

Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics*, 8(4):323–361.

Nelson, D. B. (1992). Filtering and forecasting with misspecified arch models i: Getting the right variance with the wrong model. *Journal of Econometrics*, 52(1):61–90.

Noureldin, D., Shephard, N., and Sheppard, K. (2012). Multivariate high-frequency-based volatility (heavy) models. *Journal of Applied Econometrics*, 27(6):907–933.

Pritsker, M. (2006). The hidden dangers of historical simulation. *Journal of Banking & Finance*, 30(2):561–582.

Shephard, N. and Sheppard, K. (2010). Realising the future: forecasting with high-frequency-based volatility (heavy) models. *Journal of Applied Econometrics*, 25(2):197–231.

Silvennoinen, A. and Teräsvirta, T. (2009). Multivariate garch models. In *Handbook of Financial Time Series*, pages 201–229. Springer.

Tsay, R. S. (2010). *Analysis of Financial Time Series*. Willey-Interscience, third edition.

Tse, Y. K. and Tsui, A. K. C. (2002). A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business & Economic Statistics*, 20(3):351–362.

Voev, V. and Lunde, A. (2007). Integrated covariance estimation using high-frequency data in the presence of noise. *Journal of Financial Econometrics*, 5(1):68–104.

Zhang, L., Mykland, P. A., and Aït-Sahalia, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, 100(472).

A Likelihood Estimation of Ω_t

The $(N \times N)$ random matrix S is a (centered) Wishart-matrix if it is of the form $S = XX^\top$, where the column vectors x_i of the $N \times p$ random matrix $X = (x_1, \cdots, x_p)$ are drawn from a multivariate normal distribution $x_i \sim N_N(0, \Sigma)$. The joint distribution of the matrix elements of S is called a Wishart distribution, denoted as $S \sim W_N(p, \Sigma)$. The matrix S is symmetric and positive semidefinite, since the summations of the inner product which forms the quadratic form $z^\top XX^\top z = y^\top y$ are always non negative. Furthermore, if $\Sigma > 0$ and $p > n$, then $S > 0$ (p.d.) with probability 1.

The density of S is given by

$$W_N(p, \Sigma) = \frac{|S|^{\frac{p-N-1}{2}}}{2^{\frac{Np}{2}} \Gamma_N(\frac{p}{2}) |\Sigma|^{\frac{p}{2}}} \exp(-\frac{1}{2} \operatorname{tr}(\Sigma^{-1} S)), \quad p \geq N.$$

The density for the singular Whishart distribution is

$$SINGW_N(p, \Sigma) = \frac{\pi^{\frac{-Np+p^2}{2}} \left|\widetilde{S}\right|^{\frac{p-N-1}{2}}}{2^{\frac{Np}{2}} \Gamma_N(\frac{p}{2}) |\Sigma|^{\frac{p}{2}}} \exp(-\frac{1}{2} \operatorname{tr}(\Sigma^{-1} S)), \quad p < N.$$

B Gaussian Model for R_t

$$R_t = \Omega_t^{1/2} Z_t \quad Z_t \overset{i.i.d.}{\sim} N(0, \mathbb{I}_N)$$

The model setup implies a bivariate normal distribution for the stock and market index returns R_t^i and R_t^m:

$$\left(R_t^i, R_t^m\right) \sim \mathbb{N}\left(0, \Omega_t\right) \tag{B.1}$$

By properties of the multivariate normal distribution, the marginal distributions are also normally distributed. The conditioning set for R_t^m is the $q\%$ quantile q $R_t^m < \phi^{-1}(q)\sigma_{22}$.

Applying the general law of conditional probabilities

$$\Pr(X < x | Y < y) = \frac{\Pr(X < x \wedge Y < y)}{\Pr(Y < y)},$$

to the conditional cumulative distribution function for the continuously distributed variables at hand gives:

$$p = \Pr\left(R_t^i < CoVaR | R_t^m < \phi^{-1}(q)\sigma_{22}\right) = \frac{\int_{-\infty}^{CoVaR} \int_{-\infty}^{\phi^{-1}\sigma_{22}} f_{R_t^i, R_t^m}(u, v)\, du\, dx}{q} \tag{B.2}$$

Since this expression is unwieldiy I restrain from implementing it directly and estimate the desired quantile by simulation.

C LR Test of Independce

Following Christoffersen (1998), I will apply a LR test of independence. The alternative hypothesis states that the sequence $I_t(q)$ is a Markov chain with a Markov matrix

$$\Pi_1 = \begin{pmatrix} 1-\pi_{01} & \pi_{01} \\ 1-\pi_{11} & \pi_{11} \end{pmatrix},$$

where the elements π_{ij} denote the transition probability from state i to state j: $\pi_{ij} = \Pr(I_t = j | I_{t-1} = i)$. The null hypothesis states that transition probabilities are equal and hence $\pi_{01} = \pi_{11} := \pi_0$.

The likelihood functions for the null and the alternative are

$$L(\Pi_1) = (1-\pi_{01})^{n_{00}} \pi_{01}^{n_{01}} (1-\pi_{11})^{n_{10}} \pi_{11}^{n_{11}},$$
$$L(\Pi_0) = (1-\pi_o)^{n_{00}+n_{10}} \pi_0^{n_{01}+n_{11}},$$

where n_{ij} is the number of observations with value i followed by j. The maximum likelihood estimates are given by

$$\widehat{\Pi}_1 = \begin{pmatrix} \frac{n_{00}}{n_{00}+n_{01}} & \frac{n_{01}}{n_{00}+n_{01}} \\ \frac{n_{10}}{n_{10}+n_{11}} & \frac{n_{11}}{n_{10}+n_{11}} \end{pmatrix}, \quad \widehat{\pi}_0 = \frac{n_{01}+n_{11}}{n_{00}+n_{01}+n_{10}+n_{11}}.$$

The likelihood ratio test statistic LR_I is

$$LR_I = -2\log\left[L(\widehat{\Pi}_0)/L(\widehat{\Pi}_1)\right] \sim \chi_1^2.$$

Declaration

I certify that the thesis I have presented for examination for the MSc degree of the Humboldt Universitt zu Berlin is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This thesis may not be reproduced without my prior written consent.

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.